43.1 PURPOSE:

Occasionally, a Mitochondrial DNA (mtDNA) result that was developed using a protocol that is no longer in use may need to be further evaluated for comparison purposes or to assess a match generated in CODIS.

43.2 RESPONSIBILITY:

DNA Unit Personnel authorized for Mitochondrial DNA Legacy Data Interpretation.

43.3 Comparing Sequences

43.3.1 When comparing sequences obtained from samples, only the regions with a common range will be evaluated. For example, if a partial sequence (16024-16365 and 73-284) is obtained for the evidentiary sample and a full sequence is obtained for the known sample, the comparison will be conducted on positions 16024-16365 and 73-284. In addition, sequence before and after the defined HV1 and HV2 regions will also be used for comparison purposes, provided this range is common to both samples.

43.4 Comparing Sequences with Length Heteroplasmy

43.4.1 Length heteroplasmy in HV1 most commonly arises when there is a substitution of a C for a T at position 161843. The reference type in HV1 is C_5TC_4. HV1 length heteroplasmy is not recorded in casework samples. Rather, sequences are truncated to fit the C_5TC_4 format.

43.4.2 The number of C nucleotides exhibited in samples with HV2 length heteroplasmy is highly variable and care must be exercised when making comparisons. In order for sequence concordance to be declared, a common length variant must be observed in the samples being compared. For instance, if a questioned sample has predominately 7 C residues preceding the T at position 310, but shows the presence of 8 Cs as well, it is concordant with samples containing any mixture of length variants of 7 or 8 Cs.

43.4.3 Where a sequence does not have a T at 310 and length heteroplasmy is present, it is notated 310 C and no insertions or deletions relative to rCRS at 303-315 are noted. If a sequence does not have a T at 310 and length heteroplasmy is not present, then the below guidance can be used:

Length variants in HV2 are commonly observed in the number of C residues preceding a T residue at position 310. The reference type in HV2 is C7TC5. A 309.1 C insertion in HV2 will be annotated C8TC5, and the insertions 309.1 C, 315.1 C will be annotated as C8TC6. It is possible to determine the dominant length variant in this region, as well as some or all of the minor length variants. The sequence can be notated to reflect these variants.

43.5 Interpretation of Sequence Comparisons

43.5.1 The following interpretations are available for sequence comparisons:

43.5.1.1 Exclusion:
If the samples differ at two or more positions (excluding length heteroplasmy), they can be excluded as coming from the same source or maternal lineage.

43.5.1.2 Inconclusive:

The comparison should be reported as inconclusive if the samples differ:

a. At a single position only (whether or not they share a common length variant between positions 302-310).

b. Only by not sharing a common length variant between positions 302-310 (all other positions are concordant).

Because of the possibility of undetected heteroplasmy, additional samples may be analyzed when two sequences differ by a single base. These samples may include blood, buccal swabs, and hair. Hair fragments from a known hair standard may be combined and processed as a single known sample.

43.5.1.3 Cannot Exclude:
If samples have the same sequence or are concordant, they cannot be excluded as coming from the same source or maternal lineage.

Sequence concordance is defined as having a common DNA base at every nucleotide position, including common length variants. For example, if one sample has evidence of a C and a T at a given position, and the other has a C, they share the C in common at that position and are concordant. However, if one sample has evidence of a C and a T at a given position, and the other has a G, these sequences are not concordant.

43.6 Sequence Confirmation

43.6.1 A second qualified analyst must have previously confirmed all non-mixture sequences regardless of interpretation or destination. Confirmation involves independently assembling sequences for the NC, RB, positive control, and evidentiary samples from the analysis data. Confirmed sequence range is defined as the shortest length of sequence obtained by two analysts.

Note: Mixture samples will not be interpreted for comparison purposes. A mixture is defined as 3 or more sites of point heteroplasmy.

43.7 Databases

43.7.1 The CODIS 9.0 database, containing 10,629 individuals, is used to determined match probability estimates. The database represents a broad geographic sampling from across the United States and covers the entire control region (16024-576). All profiles reflect the Scientific Working Group on DNA Analysis Methods (SWGDAM) mtDNA Interpretation Guidelines document dated 04/23/2019, which includes standards for sequence nomenclature. The database is maintained by the FBI Laboratory in collaboration with SWGDAM and is updated periodically.

43.8 Searching Profiles

43.8.1 The CODIS 9.0 database, containing 10,629 individuals, is used to search casework profiles in the forensic database. The range of the profile to be searched consists of the shortest range of sequence agreed upon by the analysts reviewing the data. Ambiguous base positions in both database samples and case work profiles are searched as an N (meaning the search includes all four bases at that position) and do not have any exclusionary value when compared with other sequences. Deletions are searched as a (-). Those database samples whose range(s) are fully included within the casework profile range are used in the database search result. Presently, this includes the regions 16024-16365 (HV1) and 73-340 (HV2). The database search result provides the number of database profiles that match the casework profile, as well as the number of profiles in the database that differ by up to five positions.

43.8.2 The number of C residues in samples with HV1 length heteroplasmy around positions 16183-16194 (see section 43.8) are not considered for comparison purposes.

43.8.3 All sequence polymorphisms in the sample searched are entered; however, all length variants at nucleotide positions 16193 and 309 are ignored in database searches of relevant concordant sequences. Hence, length variability in this region will not add any additional rarity to a database profile search.

43.9 Upper Bound Frequency Estimate

43.9.1 An upper bound 95% confidence interval can be calculated from the results of a database search in order to estimate the population frequency of a profile. The upper bound estimate is dependent on the size of the database, and these estimates may change as the database size changes.

43.9.2 The following formulae may be used to calculate the upper bound estimate:

43.9.2.1 In cases where the profile has been observed in a database:

$$p + 1.96 [(p) (1-p)/N]^{1/2}$$

p = x/N, where x is the number of times a profile has been observed in a population and N is the number of profiles in that population.

43.9.2.2 In cases where the profile has not been observed in a database:

$$1-\alpha^{1/N}$$

α is the confidence coefficient (0.05 for a 95% confidence interval), and N is the number of individuals in the population.

43.9.3 The results of these upper bound frequency estimate calculations for the three major population groups in the United States (African-American, Caucasian, and Hispanic) are included in the report. The calculations for other groups, whose total number in the database exceeds 100 people, may be included in the report where appropriate.

43.9.4 mtDNA haplotypes are inherited independently from autosomal loci. Therefore, mtDNA haplotype frequencies may be combined with autosomal STR and Y-STR frequencies to generate an overall profile frequency estimate.

43.10 References

Scientific Working Group on DNA Analysis Methods (SWGDAM). Interpretation Guidelines for Mitochondrial DNA Analysis by Forensic Testing Laboratories (2019) [Online]. Available: https://docs.wixstatic.com/ugd/4344b0_f61de6abf3b94c52b28139bff600ae98.pdf

43.11 Writing Reports:

Note: In the Testing Summary Section, a table will summarize the items and profile/regions which were amplified and sequenced and if a sample was entered into CODIS.

An example of the table:

Item #	Profile Obtained[1]					CODIS Entry
	GF	1A	1B	2A	2B	
1	Yes	Yes[2]	Yes[2]	Yes[2]	Yes[2]	Yes - CT and National

[1]GF = GlobalFiler STR DNA amplification kit

1A through 2B = mitochondrial DNA (mtDNA) hypervariable regions sequenced

[2]Using hypervariable regions 1 & 2 (HV1 & HV2) primers only

In the Conclusions Summary Section, a table will summarize the items, type of profile obtained (full or partial) and comparisons (if any).

An example of the table:

Item #	Description	Type	Victim (#2)
1	Trace material from hat	Full Profile	Excluded

Sequencing results will be found in the Appendix of the report.

An Example of the table:

Item	#1	#2	#3
Range (bp)	15998-16389	15998-16389	15998-16388
	16124 C	16124 C	16114 A
	16223 T	16223 T	16129 A
HV1	16319 A	16319 A	16213 A
			16223 T
			16278 T
			16355 T
			16362 C
Range (bp)	49-380	49-407	49-407
	73 G	73 G	73 G
	150 T	150 T	150 T
	152 C	152 C	152 C
HV2	263 G	263 G	182 T
	315.1 C	315.1 C	195 C
			198 T
			204 C
			263 G
			315.1 C

In the Conclusions Section, a table will summarize the CODIS 9.0 database search (if conducted).

An example of the table:

Group	Number of Observations	Individuals in Group	Upper Bound Frequency Estimate
African-American	0	2449	0.12% or 1 in 833
Caucasian	0	2609	0.11% or 1 in 909
Hispanic	0	2576	0.12% or 1 in 833

43.11.1 When a CODIS 9.0 database search (containing 10,629 individuals) is conducted, the report will include a database table containing the search results of the three major population groups of the United States: African-American, Caucasian, and Hispanic. The table may also contain other applicable populations where appropriate. In addition, the report will contain the search results of any other population where a matching type is found.

Note: When making a statement about the CODIS 9.0 database search, the following will be listed: the number of individuals searched, and the sequence ranges searched. See examples below.

Note: Below are examples of case scenarios. However, it is not an exhaustive list of all possible casework scenarios. Other reporting strategies (i.e. wording and or table changes) may be used with approval of the TL.

43.11.2 Cases in which the DNA sequences are the same:

43.11.2.1 Example:

The mtDNA sequences obtained from items #1 and #2 are the same. Therefore, J. DOE (or another member of the same maternal lineage) cannot be excluded as the source of item #1. Searching the mtDNA population database (CODIS 9.0, containing 10,629 individuals, at positions 16024 – 16365 and 73 – 340) the mtDNA sequence obtained from items #1 and #2 has been observed in African American, Caucasian, and Hispanic populations as follows:

[See 43.11 for an example of the CODIS 9.0 database search table]

43.11.3 Cases in which the DNA sequences are the same, but a complete sequence was not obtained from all regions:

43.11.3.1 Example:

The mtDNA sequences obtained from items #1 (at positions 16024 – 16236 and 73 – 340 only) and #2 are the same. Therefore, J. DOE (or another member of the same maternal lineage) cannot be excluded as the source of item #1.

Searching the mtDNA population database (CODIS 9.0, containing 10,629 individuals, at positions 16024 – 16236 and 73 – 340), the mtDNA sequence obtained from items #1 and #2 has been observed in African American, Caucasian, and Hispanic populations as follows:

[See 43.11 for an example of the CODIS 9.0 database search table]

Note: Above statements identifies partial range that was obtained.

43.11.4 Cases in which the DNA sequences differ at two or more nucleotide positions (excluding length heteroplasmy):

43.11.4.1 Example:

The mtDNA sequences obtained from items #1 and #2 are different. Therefore, J. DOE is excluded as the source of item #1.

43.11.5 Cases in which the DNA sequences of the evidentiary sample and one known sample are the same and the DNA sequences of the evidentiary sample and a different known sample differ at two or more nucleotide positions (excluding length heteroplasmy):

43.11.5.1 Example:

The mtDNA sequences obtained from items #1 and #2 are the same. Therefore, J. DOE (or another member of the same maternal lineage) cannot be excluded as the source of item #1.

Searching the mtDNA population database (CODIS 9.0, containing 10,629 individuals, at positions 16024 – 16365 and 73 – 340), the mtDNA sequence obtained from items #1 and #2 has been observed in African American, Caucasian, and Hispanic populations as follows:

[See 43.11 for an example of the CODIS 9.0 database search table]

The mtDNA sequence obtained from item #3 is different from the sequence obtained from item #1. Therefore, J. SMITH is excluded as the source of item #1.

43.11.6 Cases in which the DNA sequences in the evidentiary and known samples are concordant with sequence heteroplasmy present:

43.11.6.1 Example:

The mtDNA sequence obtained from item #1 is concordant with the mtDNA sequence obtained from item #2. Therefore, J. DOE (or another member of the same maternal lineage) cannot be excluded as the source of item #1.

Searching the mtDNA population database (CODIS 9.0, containing 10,629 individuals, at positions 16024 – 16365 and 73 – 340), the mtDNA sequence obtained from items #1 and #2, including all four nucleotides, A, C, G, and T, at position 16069, has been observed in African American, Caucasian, and Hispanic populations as follows:

[See 43.11 for an example of the CODIS 9.0 database search table]

43.11.7 Cases in which the DNA sequences from the evidentiary and known samples differ by a single nucleotide, and no evidence of a common nucleotide is present in the questioned and/or known sample at the position of difference:

43.11.7.1 Example:

The mtDNA sequence obtained from item #1 is the same as those obtained from items #2, #3, and #4, with the exception of position 160643. At this position, the presence of a thymine (T) was observed in item #1. In items #2, #3, and #4, a cytosine (C) was observed at position 160643. Due to the one base pair difference observed between item #1 and items #2, #3, and #4, the results are inconclusive as to whether J. DOE is excluded as the source of items #2, #3, and #4.

43.11.8 Cases in which there is a sample with an apparent mixture of more than one DNA sequence:

43.11.8.1 Example:

The mitochondrial DNA (mtDNA) sequence obtained from item #1 indicates the presence of a mixture of mtDNA from more than one individual. Therefore, no conclusions can be made for this item. No other mtDNA examinations were conducted for this item.

43.11.9 Cases in which there is a sample with an apparent mixture, but a complete sequence was not obtained from all regions:

43.11.9.1 Example:

The mtDNA sequence obtained from item #1 (at positions 16024 – 16236 and 73 – 340 only) indicates the presence of a mixture of mtDNA from more than one individual. Therefore, no conclusions can be made for this item. No other mtDNA examinations were conducted for this item.

Note: Above statement identifies partial range obtained.

43.11.10 Cases in which two evidentiary samples are the same, but (1) no known samples are available for comparison or (2) they do not match the known samples tested:

43.11.10.1 Example:

The mtDNA sequences obtained from items #1 and #2 are the same. Therefore, items #1 and #2 cannot be excluded as coming from a common source. Since no known samples were submitted, no further comparisons can be made at this time.

The mtDNA sequences obtained from items #1 and #2 are the same. Therefore, items #1 and #2 cannot be excluded as coming from a common source. The mtDNA sequences obtained from items #1 and #2 are different from item #3. Therefore, J. DOE is excluded as the source of items #1 and #2.

Optional: At the analyst's discretion, mtDNA sequencing results table and database search results table can be inserted into the report for those samples that cannot be excluded as coming from a common source.

Searching the mtDNA population database (CODIS 9.0, containing 10,629 individuals, at positions 16024 – 16365 and 73 – 340), the mtDNA sequence obtained from items #1 and #2 has been observed in African American, Caucasian, and Hispanic populations as follows:

[See 43.11 for an example of the CODIS 9.0 database search table]

43.11.11 Cases in which no sequence was obtained for a sample:

43.11.11.1 Information will be captured in the Testing Summary Table.

43.11.12 Cases in which no mtDNA examination was conducted because the remains were determined to be of non-human origin:

43.11.12.1 Example:

Dr. J. Smith examined item #1. Item #1 is not of human origin; therefore, no mitochondrial DNA examinations were conducted.

43.11.13 Cases in which the mtDNA sequence is not of the requisite quality:

43.11.13.1 Example:

The mitochondrial DNA (mtDNA) sequence obtained from item #1 was not of the requisite quality, therefore no comparisons could be made for this item.

43.11.14 Missing Persons cases:

43.11.14.1 CODIS entry will be captured in the Testing Summary Table or the Conclusions Summary Table.